MAIN POINTS

1. There is no doubt that the SFST battery developed by NHTSA is a more valuable tool than nonstandardized field sobriety tests. One simple task that demonstrates an unfair non-standard test is standing on one leg with eyes closed. The subject is told to raise his or her foot approximately six inches off the ground, keeping the foot level with the ground while keeping arms at the side without raising them. If the subject is required to perform this simple task with his or her eyes closed, the failure rate climbs to nearly 90% of all sober subjects. If the subject is required to perform this test with the eyes closed and head tilted back, the failure rate is nearly 100% of all sober subjects.
2. Standardized field sobriety tests are objective to some degree, although inter-rater reliability test-retest reliability are limited. (In statistics, inter-rater reliability, inter-rater agreement, or concordance is the degree of agreement among raters. It gives a score of how much homogeneity, or consensus, there is in the ratings given by judges. Test-retest employs the same sort of criteria, reviewing the accuracy of repeatability of the tester.)
3. The standardized tests are valuable tools only if performed correctly.
4. These standardized tests should not be admitted as substantive evidence because they are probable cause tools that sound scientific in nature and tend to confuse a jury. For example, a person can be charged and convicted of drunk driving in most states based wholly upon a test such as horizontal gaze nystagmus test, even though that test might be upwards of 90% inaccurate. (This happened in an Illinois case, People v. McKown.)

BACKGROUND

- Until the mid 1970s, police departments around the country used many different types of field sobriety tests in enforcing drunk driving laws. There was little consistency or standardization in the tests being used. Concerned over this lack of consistency, the National Highway Traffic Safety Administration (NHTSA) initiated an effort to identify the best tests for enforcement use and standardize the way they were administered and scored. NHTSA sponsored a 1977 study in which researcher were asked to identify the tests being used throughout the country and recommend a "best" test battery for further development. Out of the dozens of different tests then in use, the researchers identified three-the walk-and-turn, one-leg-stand, and horizontal gaze nystagmus tests-as the most accurate, practical, and reliable tests for enforcement purposes. A subsequent 1981 study developed a standardized set of administration and scoring principles intended to promote consistency in the use of these tests. These three tests are now known as the Standardized Field Sobriety Test Battery and form the basis of a NHTSA training program for police officers.
- While all of the tests examined were found to be "alcohol sensitive", that is, performance was affected by alcohol consumption to some degree, they were not all equally accurate.
- Some of the tests being currently used by Michigan police officers were found to be more than 80% unreliable by the 1977 researchers.

From the 1977 research:

- The horizontal gaze nystagmus test (HGN), an eye test that has been approved in Michigan in People v Berger, was found to be 77% accurate.
- The Walk and Turn test was found to be 68% accurate.
- The One Leg Stand was found to be 65% accurate.

The "accuracy" numbers have been revamped in recent validation studies, and NHTSA now claims that these tests are now claimed to be 91%, 93%, or 95% accurate.

Accuracy is misleading, however, because it does not include false arrest data. It only focuses on correct arrest decisions. Therefore, if 1,000 out 1,000 people are arrested for drunk driving based upon the fact that

have vision in one or more eyes (a prerequisite for driving), then the test is 100% accurate for detecting drunk drivers if even 1 person out of the thousand is drunk. The test remains 100% accurate if only a dozen people out of a thousand are drunk or all one-thousand are drunk.

Lawmakers should be concerned with legally protecting innocent people from false arrest.

The SFST battery had a false arrest rate of 47% in 1977. Although those numbers were lowered in later studies, this was accomplished by using higher BAC levels in roadside testing. In other words, it's easy for an officer to figure out that a person with a .17 BAC is drunk, but that does not help validate a field sobriety test for .08 BAC.

In the recent NHSTA study, "The Robustness of the Horizontal Gaze Nystagmus Test," raw data from NHTSA researchers reveals that a sober person will fail the HGN test 77% of the time. This means that the test is 77% accurate for detecting a drunk motorist, but also falsely accuses a sober person of being intoxicated 77% of the time.

When the HGN test was performed incorrectly, the false arrest rate of the HGN jumped to over 90%. The officers used to conduct this study were highly trained, highly qualified officers, so inexperience was not an issue.

Greg Kane has observed that:

Measure: metal
Care about: gun

Theory of Gun Detectors

1. Everyone with a gun sets off the metal detector.
2. That man set off the metal detector, therefore he has a gun.

Measure: coordination
Care about: high alcohol

Theory of SFSTs

1. Everyone who is drunk is uncoordinated.
2. That man is uncoordinated, therefore he is drunk.

Everyone with a gun sets off the metal detector, but most people who set off the metal detector do not have a gun.

When travelers set off a metal detector, we don't arrest and convict them of gun smuggling and terrorism. Everyone with a gun sets off the metal detector, but most people who set off the metal detector do not have a gun.

Metal detectors are highly accurate at spotting people who do have a gun, but they are **not especially good at telling the difference** between gun and no-gun. On innocent people metal detectors often give the wrong answer. They ring for keys and phones and belt buckles, for lots of stuff.